Spectral pollution and eigenvalue bounds

Spectral pollution and eigenvalue bounds

Lyonell Boulton

Revised version: August 2015

Department of Mathematics and Maxwell Institute for Mathematical Sciences.
Heriot-Watt University, Edinburgh, EH14 4AS, UK.
L.Boulton@hw.ac.uk

Abstract

The Galerkin method can fail dramatically when applied to eigenvalues in gaps of the extended essential spectrum. This phenomenon, called spectral pollution, is notoriously difficult to predict and it can occur in models from relativistic quantum mechanics, solid state physics, magnetohydrodynamics and elasticity theory. The purpose of this survey paper is two-folded. On the one hand, it describes a rigourous mathematical framework for spectral pollution. On the other hand, it gives an account on two complementary state-of-the-art Galerkin-type methods for eigenvalue computation which prevent spectral pollution completely.

1 Introduction

The Galerkin method can fail dramatically when applied to eigenvalues in gaps of the extended essential spectrum. This phenomenon, called spectral pollution, is notoriously difficult to predict and it can occur in models from relativistic quantum mechanics, solid state physics, magnetohydrodynamics, electromagnetism and elasticity theory. This survey paper has two specific purposes. On the one hand, it describes a rigourous framework for spectral pollution. On the other hand, it gives an account on two complementary state-of-the-art Galerkin-type methods for eigenvalue computation which prevent spectral pollution completely.

Introductory material is to be found in §2 and §3. The former is devoted to the basic notation around the classical Galerkin method. The latter includes a few canonical examples which illustrate the many subtleties of spectral pollution.

The main body of the text is §4-6. In §4 a generalisation of the well-known theorem by H. Weyl on the stability of the essential spectrum is closely examined. This result implies a striking fact: the spectral pollution set is stable under compact perturbations.

The text then turns to the formulation of two complementary pollution-free techniques for computation of bounds for eigenvalues. One of these techniques is related to the classical Temple-Lehmann inequality and is considered in §5. It has a local character, meaning that it just allows determination of eigenvalue bounds in the vicinity of a given parameter. These bounds are optimal in a suitable setting.

The other technique, discussed in §6, does not lead to optimal spectral bounds but it has a global character. Given any trial subspace of the domain, it always renders true information about the spectrum. Moreover, it converges under fairly general conditions.

These two approaches have recently been tested in various practical settings with successful outcomes. In order to show their implementation and range of applicability, various numerical experiments are included. These experiments are performed on two benchmark models, the two-dimensional Dirichlet Laplacian and the three-dimensional isotropic resonant cavity. They illustrate a few new features of the theory which have not been reported elsewhere. They are mostly elementary, however they may serve as a motivation for more serious investigations.

The exposition is intentionally made short and concise. It only includes the very basic aspects of both theory and applications. The material is fairly self-contained, so it must be accessible to non-specialist and PhD students in Analytical and Computational Spectral Theory. A guide for further reading is found in §7.

This survey paper began as notes from a four-weeks lecture course which I delivered at the Université de Franche-Comté Besançon in the Spring of 2012. I am duly grateful to Nabile Boussaïd and colleagues from the Laboratoire de Mathématiques for countless stimulating discussions during my visit. Financial support was provided by the Université de Franche-Comté, and the British Engineering and Physical Sciences Research Council (grant EP/I00761X/1).

2 The spectrum and the Galerkin method

The classical setting around the notions of discrete and essential spectra for self-adjoint operators, leads naturally to the framework of the Galerkin method. In this classical setting the Weyl Theorem on the stability of the essential spectrum plays a prominent role.

2.1 Nature of the spectrum for self-adjoint operators

Let $A : \operatorname{dom} A \longrightarrow \mathcal{H}$ be a densely defined self-adjoint operator on the infinite-dimensional separable Hilbert space $\mathcal{H}$. The *spectrum* of A,

$$\operatorname{spec} A = \{\lambda \in \mathbb{R} : (A - \lambda) \text{ does not have a bounded inverse}\} \quad ,$$

can be characterised by Weyl's criterion:

$$\lambda \in \operatorname{spec} A \iff \exists \{u_j\}_{j\in\mathbb{N}} \subset \operatorname{dom} A,\ \|u_j\| = 1,\ \|(A-\lambda)u_j\| \to 0 \quad .$$

The sequence of vectors $(u_j) \equiv (u_j)_{j=1}^{\infty}$ is called a *Weyl sequence (associated to λ)*. The *singular Weyl sequences* are the ones such that in addition are weakly convergent to zero[1], $u_j \rightharpoonup 0$. They determine the classical decomposition of the spectrum into two disjoint components,

$$\operatorname{spec} A = [\operatorname{spec}_{\mathrm{dsc}} A] \cup [\operatorname{spec}_{\mathrm{ess}} A] \quad .$$

The *essential spectrum* are those λ for which there is a singular Weyl sequence,

$$\lambda \in \operatorname{spec}_{\mathrm{ess}} A \quad \iff \quad \begin{cases} \exists \{u_j\}_{j\in\mathbb{N}} \subset \operatorname{dom} A, \ \|u_j\| = 1, \\ u_j \rightharpoonup 0 \ \ \& \ \ \|(A-\lambda)u_j\| \to 0 \quad . \end{cases}$$

The *discrete spectrum* is then defined as the complementary set

$$\operatorname{spec}_{\mathrm{dsc}} A = [\operatorname{spec} A] \setminus [\operatorname{spec}_{\mathrm{ess}} A] \quad .$$

The latter comprises only those $\lambda \in \mathbb{R}$ which are eigenvalues of A of finite multiplicity,

$$1 \leq \dim \ker(A - \lambda) < \infty \quad ,$$

and are isolated from the rest of the spectrum. See for example [38, §VII.3].

From the above classification of the spectrum, it is readily seen that if $B = B^*$ is another self-adjoint operator such that[2] $(A - B) \in \mathcal{K}(\mathcal{H})$, then

$$\operatorname{spec}_{\mathrm{ess}} B = \operatorname{spec}_{\mathrm{ess}} A \quad . \tag{1}$$

This observation highlights a fundamental property: the essential spectrum is a stable part of the spectrum. More generally, if A and B are *relatively compact perturbations of each other*, that is

$$(A - c)^{-1} - (B - c)^{-1} \in \mathcal{K}(\mathcal{H}) \tag{2}$$

[1] Meaning $\langle u_j, v\rangle \to 0$ for all $v \in \mathcal{H}$.

[2] Here and everywhere below $\mathcal{K}(\mathcal{H})$ is the algebra of compact operators in $\mathcal{H}$.

for at least one[3] $c \notin \mathbb{R}$, then once again (1) holds true. This stability character of the essential spectrum, and other further generalisations, are usually identified in the literature as Weyl's Theorems, see [39, Theorem XIII.14].

Weyl's Theorems as well as other classical tools in Spectral Theory such as Floquet-Bloch decompositions, allow the analytical determination of the essential spectrum for a large class of self-adjoint operators arising in applications. By contrast, only in a small handful of cases the discrete spectrum can be found explicitly. A fundamental problem in Computational Spectral Theory is the numerical estimation of eigenvalues (isolated or otherwise).

2.2 The Galerkin method

Given a linearly independent set $\{b_k\}_{k=1}^n \subset \operatorname{dom} A$ and its corresponding *trial subspace* $\mathcal{L}_n = \operatorname{span}\{b_k\}_{k=1}^n$.

a) Can we obtain rigourous spectral information about the operator A, from the action of A on $\mathcal{L}_n$?

b) If so, how can we extract this information in an optimal manner?

A partial answer to this question for semi-bounded operators is provided by the classical *Galerkin method* which is based on the Min-max Principle. See [18, Chapter 4] or [39, Theorem XIII.1].

Assume momentarily that $A = A^* \geq b > -\infty$. Let the *variational eigenvalues of* A be the non-decreasing sequence

$$\mu_k(A) = \inf_{\substack{\mathcal{V} \subset \operatorname{dom} A \\ \dim \mathcal{V} = k}} \sup_{0 \neq u \in \mathcal{V}} \frac{\langle Au, u\rangle}{\|u\|^2} = \sup_{\substack{\mathcal{V} \subset \operatorname{dom} A \\ \dim \mathcal{V}^\perp = k-1}} \inf_{0 \neq u \in \mathcal{V}} \frac{\langle Au, u\rangle}{\|u\|^2} \quad .$$

Let

$$\mu(A) = \lim_{k \to \infty} \mu_k(A) \leq \infty \quad .$$

By virtue of the Rayleigh-Ritz Theorem,

$$(-\infty, \mu(A)) \cap \operatorname{spec} A = (-\infty, \mu(A)) \cap \operatorname{spec}_{\mathrm{dsc}} A = \{\mu_k(A)\}_{k=1}^\infty \setminus \{\mu(A)\}$$

and

$$\mu(A) = \inf \operatorname{spec}_{\mathrm{ess}} A \quad .$$

Here we do not rule out the possibility of $\mu_k(A)$ being eventually constant and so $\mu(A)$ becoming an eigenvalue of infinite multiplicity, isolated or otherwise. It is routine to show that the "infimum" in the latter is a "minimum" unless $\mu(A) = \infty$, in which case A has a compact resolvent. See [39, §XIII.1].

Let the $n \times n$ hermitian matrices

$$\mathfrak{L}_n = [\langle Ab_j, b_k\rangle]_{jk=1}^n \qquad \& \qquad \mathfrak{M}_n = [\langle b_j, b_k\rangle]_{jk=1}^n \quad . \tag{3}$$

The eigenvalues of the finite-dimensional spectral problem

$$\mathfrak{L}_n \underline{u} = \lambda \mathfrak{M}_n \underline{u} \qquad 0 \neq \underline{u} \in \mathbb{C}^n \tag{4}$$

[3] Hence for all $c \notin (\operatorname{spec} A) \cup (\operatorname{spec} B)$.

are exactly the (reduced) variational eigenvalues

$$\mu_k(A, \mathcal{L}_n) = \min_{\substack{\mathcal{V} \subset \mathcal{L}_n \\ \dim \mathcal{V} = k}} \max_{0 \neq u \in \mathcal{V}} \frac{\langle Au, u \rangle}{\|u\|^2} \quad .$$

Indeed observe that the *Gram matrix* $\mathfrak{M}_n$ is positive definite and identify $u \sim \underline{u}$, where

$$u = \sum_{k=1}^{n} a_k b_k \in \mathcal{L}_n \qquad \& \qquad \underline{u} = (a_k)_{k=1}^n \in \mathbb{C}^n \quad .$$

Let $\Pi_n : \mathcal{H} \longrightarrow \mathcal{L}_n$ be the orthogonal projection $\Pi_n^2 = \Pi_n = \Pi_n^*$ onto $\mathcal{L}_n$ and let $A_n = \Pi_n A \restriction \mathcal{L}_n : \mathcal{L}_n \longrightarrow \mathcal{L}_n$ be the *reduced operator*. Then

$$\begin{aligned} \operatorname{spec} A_n &= \{\mu_1(A, \mathcal{L}_n) \leq \ldots \leq \mu_n(A, \mathcal{L}_n)\} \\ &= \{\mu_1(A_n) \leq \ldots \leq \mu_n(A_n)\} \quad . \end{aligned}$$

Therefore, as $\mu_k(A) \leq \mu_k(A, \mathcal{L}_n)$, the kth eigenvalue of the reduced operator is a guaranteed upper bound for the kth variational eigenvalues of A. Moreover, under suitable conditions, $\mu_k(A_n) \downarrow \mu_k(A)$ in the large n limit for any fixed $k \in \mathbb{N}$. These two observations form the essence of the classical Galerkin method.

The Galerkin method is of remarkable importance in the context of semi-definite operators with a compact resolvent, as it provides certified one-sided bounds for eigenvalues. The following is a canonical spectral problem which is a benchmark in this, the simplest possible setting.

The eigenvalues of the Dirichlet Laplacian on a bounded polygon Ω in $\mathbb{R}^2$ are determined by the boundary value problem

$$\begin{cases} -\Delta u = \lambda u & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \quad . \tag{5}$$

The corresponding self-adjoint operator

$$L : H_0^2(\Omega) \longrightarrow L^2(\Omega)$$

is positive definite and it has a compact (Hilbert-Schmidt) resolvent. Suppose that $\{b_k\}_{k=1}^{n(h)}$ is a basis of finite elements[4] on a regular simplicial decomposition of Ω, say of Lagrange or Hermite type, of maximum element diameter $h > 0$. Different upper estimates on the residual $\mu_k(A_h) - \mu_k(A)$ in terms of h have been extensively studied for well over 50 years. See [42, Chapter 6] and the comprehensive list of references in [17, p.283-286]. Under suitable conditions on the basis of finite elements and the regime at which the mesh becomes dense in Ω as $h \to 0$, it is guaranteed that

$$\operatorname{spec} L_h \to \operatorname{spec}_{\mathrm{dsc}} L = \operatorname{spec} L \quad .$$

[4] In the framework of the finite element method, here and elsewhere we will assume without further mention that the small real parameter h is mapped into the large integer parameter $n = n(h)$. We will only write the former as the sub-index of operators, trial subspaces, etc.

Remark 1. *For an operator A which is semi-bounded above instead of being semi-bounded below, say $A = A^* \leq b < \infty$, the Rayleigh-Ritz Theorem can be formulated for the operator $-A$ instead of A. Therefore, the Galerkin method provides reliable means of computing the eigenvalues which are outside the convex hull of the extended essential spectrum of any self-adjoint operator*[5].

The simplest case where the Galerkin method turns out to be reliable occurs when $A \in \mathcal{K}(\mathcal{H})$. In that case it is ensured that

$$\operatorname{spec} A_n \to \operatorname{spec} A$$

whenever $\Pi_n \to I$. In fact we can be more precise.

Proposition 1. *Let A be a compact self-adjoint operator. Suppose that A_n has l negative eigenvalues and m positive eigenvalues. Then*

a) A has at least l negative eigenvalues and the first l of them counting from $-\infty$ are bounded above by $\mu_k(A_n)$ for $k \in \{1, \ldots, l\}$,

b) A has at least m positive eigenvalues and the first m of them counting from $+\infty$ are bounded below by $\mu_{n+1-k}(A_n)$ for $k \in \{1, \ldots, m\}$.

Convergence of these bounds occurs as $n \to \infty$.

Proof. See [16, §5.4.2].

Note that the two statements in this proposition are not incompatible with each other, because we are counting the sequence $\operatorname{spec} A_n$ starting from its two extrema inwards.

Remark 2. *In a large number of applications involving the Galerkin method, the operator A is positive definite. In such case the* weak eigenvalue problem*,*

$$\textit{find } 0 \neq u \in \mathcal{L}_n \textit{ and } \nu > 0 \textit{ such that } \langle A^{1/2}u, A^{1/2}v\rangle = \nu\langle u, v\rangle \ \forall v \in \mathcal{L}_n \quad ,$$

is equivalent to (4)*. The formulation of the former only requires that $\mathcal{L}_n \subset \operatorname{dom} A^{1/2}$. From the Min-max principle and an argument involving the fact that $\operatorname{dom} A$ is dense in $\operatorname{dom} A^{1/2}$, it can be shown that the bounds*[6] *$\nu_k \geq \mu_k$ still hold true, if we impose this less restrictive condition on the trial subspaces. This is certainly more convenient in e.g. applications involving partial differential equations, as it means that less regularity on the trial functions is required. We focus our attention here to the more restrictive case $\mathcal{L}_n \subset \operatorname{dom} A$, because this will be required in the formulation of principles for the computation of complementary bounds for eigenvalues discussed below.*

[5] Here "extended" refers to adding topologically $+\infty$ or $-\infty$ to the essential spectrum, whenever there is accumulation of spectrum there.

[6] The ν_k here denote the eigenvalues of the weak eigenvalue problem ordered non-decreasingly.

3 Spectral pollution

The Galerkin method might not provide reliable information about the possible points of $\mathrm{spec}_{\mathrm{dsc}}\, A$ which lie inside the convex hull of the extended essential spectrum. This phenomenon can be illustrated at the practical level by means of a few striking examples.

3.1 Dichotomies in the finite section of operators

Firstly consider a very simple model from the theory of truncated Toeplitz operators [6].

Let S be the multiplication operator by the "sign" function,

$$Sf(t) = \mathrm{sign}(t) f(t) \qquad S : L^2(-\pi,\pi) \longrightarrow L^2(-\pi,\pi) \quad . \tag{6}$$

Then

$$\mathrm{spec}\, S = \mathrm{spec}_{\mathrm{ess}}\, S = \{\pm 1\} \quad .$$

The discrete Fourier transform $\mathcal{U} : L^2(-\pi,\pi) \longrightarrow \ell^2(\mathbb{Z})$

$$\mathcal{U}f = (\hat{f}(n))_{n\in\mathbb{Z}} \qquad \hat{f}(n) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(t) e^{-int} \mathrm{d}t$$

is an invertible isometry. The corresponding Laurent operator associated to S is

$$L(\mathrm{sign}) = \mathcal{U}S\mathcal{U}^{-1} = [\widehat{\mathrm{sign}}(j-k)]_{jk=-\infty}^{\infty} : \ell^2(\mathbb{Z}) \longrightarrow \ell^2(\mathbb{Z}) \quad .$$

Its spectrum coincides exactly with that of S.

Now let

$$S_n = [\widehat{\mathrm{sign}}(j-k)]_{jk=-n}^{n} \quad . \tag{7}$$

Then S_n is the reduced operator of S on

$$\mathcal{L}_n = \mathrm{span}\left\{ \frac{e^{ikt}}{\sqrt{2\pi}} \right\}_{k=-n}^{n} \subset L^2(-\pi,\pi) \quad . \tag{8}$$

Let $T(\mathrm{sign}) : \ell^2(\mathbb{N}) \longrightarrow \ell^2(\mathbb{N})$ be the Toeplitz operator

$$T(\mathrm{sign}) = [\widehat{\mathrm{sign}}(j-k)]_{jk=0}^{\infty} \quad .$$

Then S_n is also a matrix representation of $\Pi_n T(\mathrm{sign}) \restriction \mathcal{L}_n$ now on

$$\mathcal{L}_n = \{(v_k)_{k=1}^{\infty} \in \ell^2(\mathbb{N}) : v_k = 0 \ \forall k > 2n+1\} \subset \ell^2(\mathbb{N}) \quad .$$

Note that

$$\mathrm{spec}(T(\mathrm{sign})) = [-1,1]$$

[6, §1.8].

Which spectrum does S_n capture in the large n limit? Is it the one of the Laurent operator, the one of the Toeplitz operator or something in between? Since

$$\widehat{\mathrm{sign}}(j-k) = \begin{cases} 0 & k \equiv_2 j \\ \frac{2i}{(k-j)\pi} & k \not\equiv_2 j \end{cases} \quad ,$$

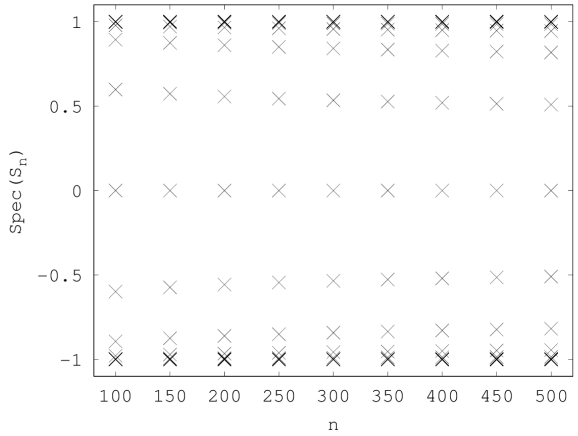


Figure 1. Computation of spec S_n for $100 \leq n \leq 500$ where $n \equiv_{50} 0$. Observe that $\operatorname{spec}(S, \mathcal{L}) = [-1, 1]$. However the eigenvalues of the reduced matrix accumulate at a much faster rate near $\operatorname{spec} S = \{\pm 1\}$. See §4.1.

then $S_n \in \mathbb{C}^{(2n+1)\times(2n+1)}$ has $n+1$ columns with odd entries equal to zero and n columns with even entries equal to zero. The former have only n non-zero entries, so they must be linearly dependent. Therefore $0 \in \operatorname{spec} S_n$ for all n, even though $0 \notin \operatorname{spec} S$.

In a situation where there are eigenvalues of the finite-dimensional problems (4) prevailing and accumulating at the resolvent set as $n \to \infty$, we say that we are in the presence of *spectral pollution*. These points of accumulation are often called *spurious eigenvalues*.

In the case of the self-adjoint operator S, for instance, $0 \notin \operatorname{spec} S$ is a spurious eigenvalue and it is by no means the only one. According to classical results [6] mainly due to Szegö, for any $\alpha \in [-1, 1]$ there exist $\alpha_n \in \operatorname{spec} S_n$ such that $\alpha_n \to \alpha$. That is, the whole segment $(-1, 1)$ fills up with spurious eigenvalues as n increases. In general, the limit of the spectrum is not the spectrum of the limit in gaps of the essential spectrum. See Figure 1.

More can be said about this very simple type of models. Standard results from the theory of truncated Toeplitz matrices [6] provide a satisfactory explanation of the appearance of spectral pollution in this particular case. However, from a general perspective, it shows that the Galerkin method can fail dramatically when applied in gaps of the essential spectrum. If we did not know the spectrum beforehand, from the above analysis we might be driven to a false conclusion that 0, for instance, is (or is near) it.

3.2 Indefinite operators

We might get misleading information about possible points in the spectrum inside the convex hull of the extended essential spectrum, even when

the operator has a compact resolvent and the trial spaces comprise only finite combinations of eigenvectors. In order to see how this can occur in a concrete setting, consider the model of a resonant electromagnetic cavity with perfect conductivity through the boundary.

Let the convex polyhedron $\Omega \subset \mathbb{R}^3$ be filled with an isotropic medium (dimensionless, unit electric permittivity and magnetic permeability). The physical phenomenon of electromagnetic oscillations in Ω is described by the time independent Maxwell system

$$\begin{cases} \operatorname{curl} \underline{\mathrm{E}} = i\omega \underline{\mathrm{H}} & \text{in } \Omega \\ \operatorname{curl} \underline{\mathrm{H}} = -i\omega \underline{\mathrm{E}} & \text{in } \Omega \\ \underline{\mathrm{E}} \times \underline{\mathrm{n}} = 0 & \text{on } \partial\Omega \end{cases} \tag{9}$$

for unknown angular frequencies $\omega \in \mathbb{R}$ and non-zero solenoidal field phasors $[\underline{\mathrm{E}}, \underline{\mathrm{H}}]^{\mathrm{t}} \in \mathcal{J}$, see (10). The normal unit vector on $\partial\Omega$ is written as $\underline{\mathrm{n}}$.

The self-adjoint operator $\mathcal{M} : \operatorname{dom} \mathcal{M} \longrightarrow \mathcal{H}$ associated to (9) is [4]

$$\underbrace{\begin{bmatrix} 0 & i\operatorname{curl} \\ -i\operatorname{curl} & 0 \end{bmatrix}}_{\mathcal{M}} : \underbrace{\begin{matrix} H_0(\operatorname{curl};\Omega) \\ \oplus \\ H(\operatorname{curl};\Omega) \end{matrix}}_{\operatorname{dom}\mathcal{M}} \longrightarrow \underbrace{\begin{matrix} [L^2(\Omega)]^3 \\ \oplus \\ [L^2(\Omega)]^3 \end{matrix}}_{\mathcal{H}} .$$

Here

$$H(\operatorname{curl};\Omega) = \left\{ \underline{\mathrm{F}} \in [L^2(\Omega)]^3 : \operatorname{curl} \underline{\mathrm{F}} \in [L^2(\Omega)]^3 \right\}$$

is the maximal domain of the "curl" and

$$H_0(\operatorname{curl};\Omega) = \left\{ \underline{\mathrm{F}} \in H(\operatorname{curl};\Omega) : \int_\Omega \operatorname{curl} \underline{\mathrm{F}} \cdot \underline{\mathrm{G}} = \int_\Omega \underline{\mathrm{F}} \cdot \operatorname{curl} \underline{\mathrm{G}} \quad \forall \underline{\mathrm{G}} \in H(\operatorname{curl};\Omega) \right\}$$

is the minimal domain which encodes the boundary conditions.

The *solenoidal space*

$$\mathcal{J} = \left\{ \begin{bmatrix} \underline{\mathrm{F}} \\ \underline{\mathrm{G}} \end{bmatrix} \in \operatorname{dom} \mathcal{M} : \operatorname{div} \underline{\mathrm{F}} = 0 = \operatorname{div} \underline{\mathrm{G}} \quad \& \quad (\underline{\mathrm{G}} \cdot \underline{\mathrm{n}}) \restriction_{\partial\Omega} = 0 \right\} \tag{10}$$

is compactly embedded into $[L^2(\Omega)]^6$ and it exactly coincides with

$$(\ker \mathcal{M})^\perp \cap \operatorname{dom} \mathcal{M} \quad .$$

Then $\operatorname{spec}_{\mathrm{ess}} \mathcal{M} = \{0\}$ while $\operatorname{spec}_{\mathrm{dsc}} \mathcal{M}$ is an infinite set of eigenvalues which only accumulates at $\pm\infty$. The latter is symmetric with respect to 0, because $\begin{bmatrix} \underline{\mathrm{E}} \\ \underline{\mathrm{H}} \end{bmatrix} \neq 0$ is an eigenvector of (9) associated to ω if and only if $\begin{bmatrix} -\underline{\mathrm{E}} \\ \underline{\mathrm{H}} \end{bmatrix} \neq 0$ is an eigenvector of (9) associated to $-\omega$.

The reduced operator

$$\tilde{\mathcal{M}} = \mathcal{M} \restriction \mathcal{J} : \mathcal{J} \longrightarrow [L^2(\Omega)]^6 \ominus \ker \mathcal{M}$$

has a compact resolvent and it is the one describing the physical phenomenon of electromagnetic oscillations in the isotropic resonant cavity.

This operator is strongly indefinite, and its spectrum and eigenspaces coincide exactly with those of $\mathcal{M}$ except for $\omega = 0$. The latter is not an eigenvalue of $\tilde{\mathcal{M}}$.

Example 1. Let

$$\operatorname{spec} \tilde{\mathcal{M}} = \{\pm\omega_j\}_{j\in\mathbb{N}}$$

where the positive eigenvalues are $0 < \omega_j \leq \omega_{j+1} \uparrow \infty$. Here we count the multiplicity. Let $\{\Phi_j^{\pm}\}_{j\in\mathbb{N}} \subset \mathcal{J}$ be an associated orthonormal basis of eigenfunctions

$$\tilde{\mathcal{M}}\Phi_j^{\pm} = \pm\omega_j\Phi_j^{\pm} \quad .$$

Then

$$\tilde{\mathcal{M}} = \sum_{j=1}^{\infty} \omega_j(|\Phi_j^+\rangle\langle\Phi_j^+| - |\Phi_j^-\rangle\langle\Phi_j^-|) \quad .$$

If we assume that the trial spaces are

$$\mathcal{L}_n = \operatorname{span}\left\{\Phi_1^{\pm}, \ldots, \Phi_{n-1}^{\pm}, \frac{1}{\sqrt{2}}\Phi_n^+ + \frac{1}{\sqrt{2}}\Phi_n^-\right\} \quad , \tag{11}$$

which might seem to be extremely close to the actual spectral subspaces of $\tilde{\mathcal{M}}$, it turns out that

$$\operatorname{spec} \tilde{\mathcal{M}}_n = \{0, \pm\omega_1, \ldots, \pm\omega_{n-1}\} \quad .$$

Therefore we might be falsely led to believe that $0 \in \operatorname{spec} \tilde{\mathcal{M}}$.

Remarkably, the Galerkin method successfully applies to $R = \tilde{\mathcal{M}}^{-1}$ in this case, as it is a compact operator. Picking the same trial spaces (11), we get

$$\operatorname{spec} R_n = \{\pm\omega_1^{-1}, \ldots, \pm\omega_{n-1}^{-1}, 0\} \to \operatorname{spec} R \quad ,$$

and the properties *a)* and *b)* from Proposition 1 are fulfilled.

Example 1 can be modified in order to create a dense set of spurious eigenvalues for $\tilde{\mathcal{M}}$. See [14, Example 1.2], [30] and [31]. Arguably, this example is rather artificial as the family of trial subspaces has been taylor made to generate spectral pollution. However, as we shall see next, spectral pollution also occurs in the canonical discretisation of (9) by means of the finite element method.

3.3 The finite element method and spectral pollution

The operator $\mathcal{S} = \tilde{\mathcal{M}}^2$ does not yield Galerkin spurious eigenvalues as it is semi-definite and it has a compact resolvent. For the trial spaces (11), we get

$$\operatorname{spec} \mathcal{S}_n = \{\underbrace{\omega_1^2, \ldots, \omega_{n-1}^2}_{\text{multiplicity 2}}, \omega_n^2\} \to \operatorname{spec} \mathcal{S} \quad .$$

In fact note that $\mathcal{S}$ is equivalent to a vector-valued Laplacian on Ω with suitable boundary conditions, so we are in a situation very similar to that of the Dirichlet Laplacian (5).

The operator $\mathcal{M}^2$ is completely different in this respect. All its spectrum is trapped in a gap of the extended essential spectrum, the segment

$(0,\infty)$. This has an important consequence for the numerical estimation of the angular frequencies in the resonant cavity by means of the finite element method as we illustrate next.

Let $\Omega = [0,\pi]^3$. The non-zero eigenfrequencies of $\mathcal{S}$ and $\mathcal{M}^2$ are

$$\omega = \pm\sqrt{j^2+k^2+l^2}$$

for indices $\{j,k,l\} \subset \mathbb{N}\cup\{0\}$ not two of them vanishing simultaneously. The corresponding $\underline{\mathrm{E}}$ component of the field phasors are

$$\underline{\mathrm{E}}(x,y,z) = \begin{bmatrix} \alpha\cos(jx)\sin(ky)\sin(lz) \\ \beta\sin(jx)\cos(ky)\sin(lz) \\ \gamma\sin(jx)\sin(ky)\cos(lz) \end{bmatrix} \qquad \forall \begin{bmatrix}\alpha\\ \beta\\ \gamma\end{bmatrix} \cdot \begin{bmatrix} j\\ k\\ l\end{bmatrix} = 0 \quad .$$

Suppose that we did not know an analytic expression for $\operatorname{spec}\tilde{\mathcal{M}}$ and that we wanted to approximate a few eigenvalues numerically. Generating $\mathcal{L}_n$ by means of the finite element method is highly non-trivial. For example the standard nodal elements are not typically solenoidal and most of the edge elements are so, only in the interior of the simplexes but not across their boundaries. This difficulty has been well documented in the literature. See [1], [5] and references therein.

Example 2. In Figure 2 we have naïvely picked $\mathcal{L}_h \subset \operatorname{dom}\mathcal{M} \setminus \operatorname{dom}\tilde{\mathcal{M}}$ made of Lagrange elements of order 3 for the unstructured uniform mesh $\mathcal{T}_h$ depicted in (a),

$$\begin{aligned} \mathcal{V}_h &= \{\underline{\mathrm{F}} \in [C^0(\overline{\Omega})]^3 : \underline{\mathrm{F}}|_K \in [\mathbb{P}_3(K)]^3 \ \forall K \in \mathcal{T}_h\} \\ \mathcal{V}_{h,0} &= \{\underline{\mathrm{G}} \in \mathcal{V}_h : (\underline{\mathrm{G}} \times \underline{\mathrm{n}}) \restriction_{\partial\Omega} = 0\} \\ \mathcal{L}_h &= \mathcal{V}_{h,0} \times \mathcal{V}_h \subset \operatorname{dom}\mathcal{M} \quad . \end{aligned} \tag{12}$$

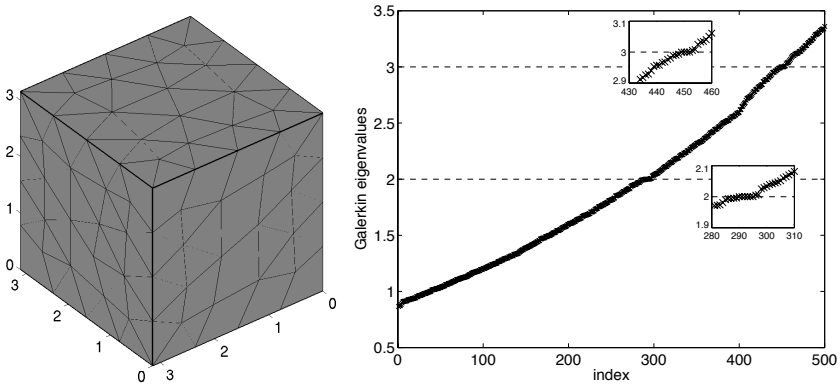
The graph (b) shows 500 eigenvalues of $(\mathcal{M}^2)_n$ near 2. The only true eigenvalues of $\mathcal{M}^2$ in the segment $[0.5, 3.5]$ are $\omega_1^2 = 2$ of multiplicity 3 and $\omega_2^2 = 3$ of multiplicity 2. The picture does not give much information about the true spectrum of the operator in this segment.

Of course we know analytically what the eigenvalues are in this example. However, if we did not know the spectrum[7], then a direct application of the Galerkin method is likely to be completely useless.

4 A rigourous framework for spectral pollution

It is possible to characterise spectral pollution by means of particular classes of Weyl sequences. This allows the formulation of suitable versions of Weyl-type theorems in the context of eigenvalue computation.

[7] Such is the case when Ω has a more complicated shape.



(a) $\dim \mathcal{L}_{0.9} \approx 22\,\mathrm{K}$ **(b)** 500 eigenvalues of $(\mathcal{M}^2)_n$ nearest to 2

Figure 2. The only eigenvalues of $\mathcal{M}^2$ in the segment $[0.5, 3.5]$ are $\omega^2 = 2$ of multiplicity 3 and $\omega^2 = 3$ of multiplicity 2. However the implementation of the Galerkin method shown in (b) wrongly suggests that the segment $[0.9, 3.4]$ is in the spectrum. The subspace $\mathcal{L}_n$ is made of Lagrange elements of order 3 in the mesh shown in (a).

4.1 The limit of the spectra vs the spectrum of the limit

The following is an elementary property for self-adjoint operators and it turns out to play a fundamental role below.

Lemma 2. *Let B be a self-adjoint operator. The Hausdorff distance from any $t \in \mathbb{R}$ to the spectrum of B can be determined via*

$$\operatorname{dist}(t, \operatorname{spec} B) = \inf_{0 \neq u \in \operatorname{dom} B} \frac{\|(B-t)u\|}{\|u\|} \quad . \tag{13}$$

Proof. Indeed

$$\begin{aligned} \inf_{0 \neq u \in \operatorname{dom} B} \frac{\|(B-t)u\|^2}{\|u\|^2} &= \min\{\lambda \in \operatorname{spec}(B-t)^2\} \\ &= \operatorname{dist}\left(0, \operatorname{spec}(B-t)^2\right) \quad . \end{aligned}$$

By taking square roots, as appropriate, (13) follows.

Let $\mathcal{L} = \{\mathcal{L}_n\}_{n \in \mathbb{N}}$ be a family of trial subspaces $\mathcal{L}_n \subset \operatorname{dom} A$. We say that $\mathcal{L}$ is *regular*, if for any $u \in \operatorname{dom} A$ we can find $u_n \in \mathcal{L}_n$ such that

$$\|u - u_n\| + \|Au - Au_n\| \to 0 \quad .$$

Denote the *outer limit spectrum of the Galerkin method* (applied to A) with respect to the sequence $\mathcal{L}$ by

$$\operatorname{spec}(A, \mathcal{L}) = \limsup_{n \to \infty} \operatorname{spec} A_n = \{\lambda \in \mathbb{R} : \exists \lambda_j \in \operatorname{spec} A_{n_j},\, \lambda_j \to \lambda\} \quad .$$

What would the relation between this upper limit and the true spectrum of A be?

For any given $\lambda \in \operatorname{spec} A$ and non-zero $u \in \operatorname{dom} A$ such that

$$\frac{\|(A-\lambda)u\|}{\|u\|} < \frac{\varepsilon}{2} \quad ,$$

the condition of regularity on $\mathcal{L}$ implies that for n large

$$\frac{\|\Pi_n(A-\lambda)\Pi_n u\|}{\|\Pi_n u\|} < \varepsilon \quad .$$

Then, according to (13) for $B = A_n$, there exists $\lambda_n \in \operatorname{spec} A_n$ such that $|\lambda - \lambda_n| < \varepsilon$ in the large n regime. Hence, taking $\varepsilon \to 0$, it follows that[8]

$$\operatorname{spec} A \subseteq \operatorname{spec}(A, \mathcal{L}) \quad .$$

In §3 we showed with a concrete example that we must not expect an equality here in general. The difference between the two sets is the *region of spectral pollution* for the Galerkin method.

Example 3. Let S and $\mathcal{L}$ be as in (6) and (8). Let $e_0(x) = \frac{1}{\sqrt{2\pi}}$. Set $K = |e_0\rangle\langle e_0|$ and $B = S + K$. Then [7, Lemma 7]

$$\operatorname{spec} B = \underbrace{\{\pm 1\}}_{\operatorname{spec}_{\mathrm{ess}} B} \cup \underbrace{\{(1 \pm \sqrt{5})/2\}}_{\operatorname{spec}_{\mathrm{dsc}} B}$$

and the eigenvalues in the discrete spectrum are both simple. By virtue of Theorem 4 below,

$$\operatorname{spec}(B, \mathcal{L}) = [-1, 1] \cup \{(1+\sqrt{5})/2\} \quad . \tag{14}$$

See Figure 3.

Remark 3. *Despite of the spectral pollution phenomenon illustrated in Figure 3, note that the eigenvalues of B_n accumulate much faster at the spectrum of B than at any other point of the segment $(-1, 1)$.*

4.2 Singular Weyl sequences

The appearance of spurious eigenvalues can be characterised in a fairly general context by the existence of particular Weyl sequences of singular type.

Lemma 3. *$\lambda \in \operatorname{spec}(A, \mathcal{L})$ if and only if there exists a sequence $\{v_j\}_{j\in\mathbb{N}} \subset \operatorname{dom} A$ such that $v_j \in \mathcal{L}_{n_j}$ with $\|v_j\| = 1$ and*

$$\lim_{k\to\infty} \|\Pi_{n_j}(A-\lambda)\, v_j\| = 0 \quad .$$

Proof. According to the definition, $\lambda \in \operatorname{spec}(A, \mathcal{L})$ if and only if there exists $\lambda_j \in \mathbb{R}$ and $v_j \in \mathcal{L}_{n_j}$ with $\|v_j\| = 1$ such that $\lambda_j \to \lambda$ and $\Pi_{n_j}(A - \lambda_j)\, v_j = 0$. As

$$\Pi_{n_j}(A-\lambda)\, v_j = (\lambda_j - \lambda) v_j \to 0 \quad ,$$

[8] This well-known inclusion has a long history. See [36].

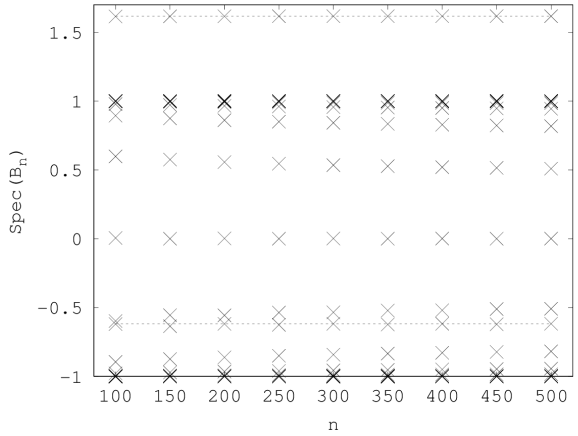


Figure 3. Spectrum of the reduced operator B_n in Example 3 for $100 \leq n \leq 500$ where $n \equiv_{50} 0$. In the picture the two points in the discrete spectrum $(1 \pm \sqrt{5})/2$ have been highlighted. See (14).

one of the stated implications follows immediately.

On the other hand, if $\{v_j\}_{j\in\mathbb{N}} \subset \operatorname{dom} A$ is as stated, then $\|(A_{n_j} - \lambda)v_j\| \to 0$. Since the A_n are hermitian, by virtue of Lemma 2 there exists $\lambda_j \in \operatorname{spec} A_{n_j}$ such that

$$|\lambda_j - \lambda| \leq \|(A_{n_j} - \lambda)v_j\| \to 0 \quad .$$

Thus $\lambda \in \operatorname{spec}(A, \mathcal{L})$ ensuring the complementary implication.

A sequence (v_j) satisfying the condition of Lemma 3 will be called an *$\mathcal{L}$-Weyl sequence* for $\lambda \in \operatorname{spec}(A, \mathcal{L})$. If additionally $v_j \nrightarrow 0$, then it will be called a *singular $\mathcal{L}$-Weyl sequence*.

By analogy to the classical notions, we call

$$\operatorname{spec}_{\text{ess}}(A, \mathcal{L}) = \{\lambda \in \mathbb{R} : \exists \text{ a singular } \mathcal{L}\text{-Weyl sequence for } \lambda\}$$

the *limit essential spectrum of the Galerkin method* applied to A with respect to the sequence $\mathcal{L}$. By virtue of Lemma 3, $\operatorname{spec}_{\text{ess}}(A, \mathcal{L}) \subseteq \operatorname{spec}(A, \mathcal{L})$. Moreover,

$$\operatorname{spec}_{\text{ess}} A \subseteq \operatorname{spec}_{\text{ess}}(A, \mathcal{L}) \quad .$$

The following statement is an immediate consequence of the above definition.

Theorem 4 (Version of Weyl's Theorem for spectral pollution)**.** *Let $\mathcal{L}$ be a regular family associated to the self-adjoint operator A. Let K be a compact self-adjoint operator. Then*

$$\operatorname{spec}_{\text{ess}}(A, \mathcal{L}) = \operatorname{spec}_{\text{ess}}(A + K, \mathcal{L}) \quad . \tag{15}$$

The identity (15) is still holds true under much weaker conditions on the perturbation K. However, at present, it is unclear whether these match (2) in full generality. See [10] for details.

As we see next, spectral pollution only occurs in the limit essential spectrum of the Galerkin method. Therefore, by combining the classical Weyl's Theorem with Theorem 4, the closure of the region of spectral pollution is unchanged under compact perturbations. The *limit discrete spectrum of the Galerkin method* is the set

$$\operatorname{spec}_{\mathrm{dsc}}(A,\mathcal{L}) = \operatorname{spec}(A,\mathcal{L}) \setminus \operatorname{spec}_{\mathrm{ess}}(A,\mathcal{L}) \quad .$$

Lemma 5. *Let (v_j) be an $\mathcal{L}$-Weyl sequence for $\lambda \in \operatorname{spec}(A,\mathcal{L})$. If $v_j \rightharpoonup v$, then $v \in \ker(A-\lambda)$.*

Proof. Let $f \in \operatorname{dom} A$. Let $f_n \in \mathcal{L}_n$ be such that $f_n \to f$ and $Af_n \to Af$. Then

$$\begin{aligned}\langle (A_{n_j}-\lambda)v_j, f\rangle &= \langle \Pi_{n_j}(A-\lambda)v_j, f\rangle \\ &= \langle v_j, (A-\lambda)f_{n_j}\rangle \to \langle v, (A-\lambda)f\rangle \quad .\end{aligned}$$

Hence

$$\langle (A-\lambda)v, f\rangle = 0 \qquad \forall f \in \operatorname{dom} A$$

and so $(A-\lambda)v = 0$.

By virtue of this lemma,

$$\operatorname{spec}_{\mathrm{dsc}}(A,\mathcal{L}) \subseteq \operatorname{spec}_{\mathrm{dsc}} A \tag{16}$$

and therefore spectral pollution can only appear in $\operatorname{spec}_{\mathrm{ess}}(A,\mathcal{L})$.

The present framework for spectral pollution is based on a natural separation of the upper limit spectrum of the Galerkin method into a discrete part and an essential part. As we shall see from the next statement, it also allows a natural classification of the spurious eigenvalues. A complete proof and further details on the matter can be found in [10].

Theorem 6 (Nature of the spectral pollution phenomenon)**.** *Let $\mathcal{L}$ be a regular family associated to the self-adjoint operator A. Then $\lambda \in \operatorname{spec}_{\mathrm{ess}}(A,\mathcal{L})$ if and only if one and only one of the following possibilities holds true.*

a) $\lambda \notin \operatorname{spec} A$, and there exist $\lambda_j \to \lambda$ and $v_j \in \mathcal{L}_{n_j}$ such that

$$\|v_j\| = 1, \quad A_{n_j}v_j = \lambda_j v_j \quad \& \quad v_j \rightharpoonup 0 \quad .$$

b) $\lambda \in \operatorname{spec}_{\mathrm{ess}} A$, and there exist $\lambda_j \to \lambda$ and $v_j \in \mathcal{L}_{n_j}$ such that

$$\|v_j\| = 1, \quad A_{n_j}v_j = \lambda_j v_j \quad \& \quad v_j \rightharpoonup 0 \quad .$$

c) $\lambda \in \operatorname{spec}_{\mathrm{dsc}} A$ and for all $\varepsilon > 0$

$$\#\{\lambda \in \operatorname{spec}(A_{n_j}) \cap (\lambda-\varepsilon, \lambda+\varepsilon)\} > \dim\ker(A-\lambda) \quad .$$

Case *a)* means that λ is in the region of spectral pollution, while in cases *b)* and *c)* λ is in the spectrum. For *b)* we can say nothing more, but for *c)* the multiplicity of λ is wrongly predicted by $\mathcal{L}$.

Example 4. Let $\mathcal{L}$ and B be as in Example 3. By virtue of (16),

$$\operatorname{spec}_{\text{ess}}(B,\mathcal{L}) = [-1,1] \qquad \& \qquad \operatorname{spec}_{\text{dsc}}(B,\mathcal{L}) = \{(1+\sqrt{5})/2\} \quad .$$

In the context of Theorem 6:

a) holds in $(-1,1) \setminus \{(1-\sqrt{5})/2\}$,

b) holds at $\lambda = \pm 1$ and

c) holds at $\lambda = (1-\sqrt{5})/2$.

Even in situations where the Galerkin method converges to gap eigenvalues and there is no spectral pollution, there is no guarantee of a monotone convergence in general. Recall the question *a)* in §2.2. We now turn to examine two Galerkin-type approaches for computation of guaranteed bounds for eigenvalues which work in any part of the spectrum.

5 Local bounds for eigenvalues

A family of computational methods which prevent spectral pollution is closely linked with generalisations [47, Chapter 4] of the classical Temple-Lehmann bound [18, p.93] for eigenvalues of semi-definite self-adjoint operators. These methods have a *local* character. They only determine information about the spectrum in the vicinity of a given point $t \in \mathbb{R}$ which is set beforehand.

The choice of t is normally a delicate issue. It might require some analytical information about the rough position of the spectrum to start with. Homotopy methods [34, 35] are among the best approaches to address this issue. When applicable, these local methods tend to be more accurate than a "global" counterpart examined in the next section.

5.1 Approximated spectral distances

For $z \in \mathbb{C}$, let

$$F_n(z) = \min_{0 \neq u \in \mathcal{L}_n} \frac{\|(A-z)u\|}{\|u\|} \quad .$$

By virtue of Lemma 2,

$$F_n(z) \geq \operatorname{dist}(z, \operatorname{spec} A) \quad . \tag{17}$$

Hence

$$[t - F_n(t), t + F_n(t)] \cap \operatorname{spec} A \neq \varnothing \qquad \forall t \in \mathbb{R} \quad . \tag{18}$$

Moreover, if $\mathcal{L} = \{\mathcal{L}_n\}_{n\in\mathbb{N}}$ is a regular family, then

$$\lim_{n\to\infty} F_n(t) = \operatorname{dist}(t, \operatorname{spec} A) \quad .$$

Therefore $F_n(t)$ is an upper approximation of the distance to the spectrum, under natural conditions of convergence. Here and everywhere below we will regard the parameter t as real and the parameter z as complex[9].

[9] This distinction will be needed in §6.

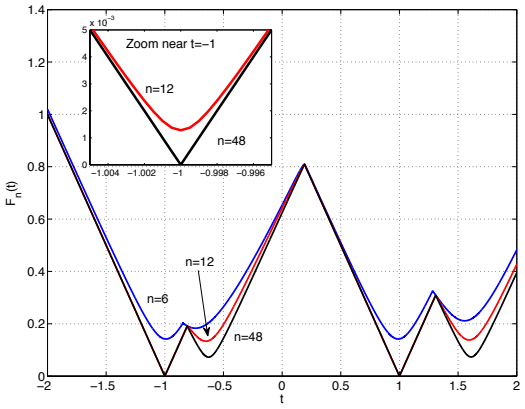


Figure 4. $F_n(t)$ for $n \in \{6, 12, 48\}$. The operator and regular family are those from Example 3. The picture suggest a better way of getting enclosures for the eigenvalues than finding the local minima of $F_n(t)$. See Proposition 8.

Let

$$\mathfrak{K}_n = [\langle Ab_j, Ab_k \rangle]_{jk=1}^n. \tag{19}$$

Let

$$\mathfrak{Q}_n(z) = \mathfrak{K}_n - 2z\mathfrak{L}_n + z^2\mathfrak{M}_n \quad \& \quad \tilde{\mathfrak{Q}}_n(z) = \mathfrak{M}_n^{-1/2}\mathfrak{Q}_n(z)\mathfrak{M}_n^{-1/2} \quad , \tag{20}$$

where $\mathfrak{L}_n$ and $\mathfrak{M}_n$ are as in (3). The computation of $F_n(t)$ can be performed via the following characterisation.

Lemma 7. *For $t \in \mathbb{R}$,*

$$F_n(t)^2 = \min \operatorname{spec} \tilde{\mathfrak{Q}}_n(t) \quad .$$

Proof. Observe that

$$F_n(t)^2 = \min_{0 \neq u \in \mathcal{L}_n} \frac{\langle (A-t)u, (A-t)u \rangle}{\langle u, u \rangle} \quad .$$

According to the Rayleigh-Ritz Theorem, the right hand side is the smallest eigenvalue $\gamma \geq 0$ of the problem

$$\mathfrak{Q}_n(t)\underline{u} = \gamma \mathfrak{M}_n \underline{u} \quad .$$

This ensures the claimed statement.

We might expect that, in practice, we should find local minima of $F(t)$ to get good enclosures for point in the spectrum. In fact, there is a better way [20, 19]. By virtue of the triangle inequality,

$$F_n(t) \leq F_n(s) + |s - t| \qquad \forall s \in \mathbb{R} \quad .$$

Hence

$$|F_n(t) - F_n(s)| \leq |s - t| \qquad \forall t, s \in \mathbb{R} \quad . \tag{21}$$

That is, $F_n(t)$ is a Lipschitz continuous function with modulus of uniform continuity less than or equal to one. Let

$$\mathfrak{n}^-(t) = \sup\{\operatorname{spec} A \cap (-\infty, t]\} \quad \& \quad \mathfrak{n}^+(t) = \inf\{\operatorname{spec} A \cap [t, +\infty)\} \quad .$$

These are the points from spec A which are nearest to t, so that

$$\operatorname{dist}(t, \operatorname{spec} A) = \min\{t - \mathfrak{n}^-(t), \mathfrak{n}^+(t) - t\} \quad .$$

The next crucial observation suggests an optimal setting for extracting information about spec A from the profile of $F_n(t)$. See Figure 4.

Proposition 8. *Let $t^- < t < t^+$. Then*

$$\begin{array}{lll} F_n(t^-) \leq t - t^- & \Rightarrow & t^- - F_n(t^-) \leq \mathfrak{n}^-(t) \\ F_n(t^+) \leq t^+ - t & \Rightarrow & t^+ + F_n(t^+) \geq \mathfrak{n}^+(t) \end{array} \quad . \tag{22}$$

Moreover, let $t_1^- < t_2^- < t < t_2^+ < t_1^+$. Then

$$\begin{array}{lll} F_n(t_j^-) \leq t - t_j^- \text{ for } j = 1, 2 & \Rightarrow & t_1^- - F_n(t_1^-) \leq t_2^- - F_n(t_2^-) \leq \mathfrak{n}^-(t) \\ F_n(t_j^+) \leq t_j^+ - t \text{ for } j = 1, 2 & \Rightarrow & t_1^+ + F_n(t_1^+) \geq t_2^+ + F_n(t_2^+) \geq \mathfrak{n}^+(t) \end{array} \quad .$$

Proof. If $t \geq F_n(t^-) + t^-$, then

$$[t^- - F_n(t^-), t] \cap \operatorname{spec} A \neq \varnothing$$

and so $\mathfrak{n}^-(t) \in [t^- - F_n(t^-), t]$. The other statement in (22) is shown in a similar fashion.

By virtue of (21), the maps $t \mapsto t \pm F_n(t)$ are monotonically increasing. This ensures the second assertion.

This proposition leads to a remarkable conclusions, examined at length in [19] and [20].

a) If $t_\infty^- < t$ is such that

$$t_\infty^- + F_n(t_\infty^-) = t \quad ,$$

then $t_\infty^- - F_n(t_\infty^-)$ is a certain lower bound of $\mathfrak{n}^-(t)$ and it is optimal in the sense that any other $s < t_\infty^-$ would give $s - F_n(s) \leq t_\infty^- - F_n(t_\infty^-)$.

b) If $t_\infty^+ > t$ is such that

$$t_\infty^+ - F_n(t_\infty^+) = t \quad ,$$

then $t_\infty^+ + F_n(t_\infty^+)$ is a certain upper bound of $\mathfrak{n}^+(t)$ an it is optimal in the sense that any other $s > t_\infty^+$ would give $t_\infty^+ + F_n(t_\infty^+) \leq s + F_n(s)$.

The optimal $t_\infty^\pm$ are near the local maxima of $F_n(t)$. A hierarchical strategy for finding bounds for eigenvalues based on this simple observation was first described in [19] and then refined in [20, 21]. This strategy is equivalent to a generalisation [48] of the classical Temple-Lehmann-Goerisch method, which we describe next and which seems better suited for concrete implementations.

5.2 A realisation of the method

Fix $t \in \mathbb{R}$. The optimal parameters $t_\infty^\pm$ from *a)-b)* above can be found as follows. Let

$$\mathfrak{R}_n(t) = \mathfrak{L}_n - t\mathfrak{M}_n$$

and consider the t-dependent matrix eigenvalue problem

$$\mathfrak{R}_n(t)\underline{u} = \tau\mathfrak{Q}_n(t)\underline{u} \quad . \tag{23}$$

Both the eigenvalues $\tau \equiv \tau(t)$ and eigenvectors $\underline{u} \equiv \underline{u}(t)$ here depend on t. As this is always clear from the context below, we do not highlight this dependence explicitly.

Proposition 9. *If $\tau \neq 0$ is an eigenvalue of (23), then $F_n\left(t + \frac{1}{2\tau}\right) \leq \frac{1}{2|\tau|}$. Moreover, if the identity*

$$F_n\left(t + s\right) = |s|$$

is satisfied for some $s \neq 0$, then $\tau = \frac{1}{2s}$ is an eigenvalue of (23).

Proof. Recall Lemma 7 and note that

$$\mathfrak{Q}_n\left(t + s\right)\underline{u} = s^2\mathfrak{M}_n\underline{u}$$

can be rewritten as (23) with the substitution $s = \frac{1}{2\tau}$.

We denote the smallest and the largest eigenvalue of (23) by τ_- and τ_+, respectively. That is

$$\begin{aligned}\tau_- &= \min\{\tau \in \mathbb{R} : \det[\mathfrak{R}_n(t) - \tau\mathfrak{Q}_n(t)] = 0\} \\ \tau_+ &= \max\{\tau \in \mathbb{R} : \det[\mathfrak{R}_n(t) - \tau\mathfrak{Q}_n(t)] = 0\} \quad .\end{aligned}$$

Theorem 10 (Local bounds for eigenvalues)**.** *Let $t \in \mathbb{R}$.*

a) If $\tau_- < 0$, then $t + \frac{1}{\tau_-} \leq \mathfrak{n}^-(t)$.

b) If $\tau_+ > 0$, then $t + \frac{1}{\tau_+} \geq \mathfrak{n}^+(t)$.

Proof. The proof of both statements is similar. Consider that of *a)*. Let $t^- = t + \frac{1}{2\tau_-}$. By Proposition 9, $F_n(t^-) \leq t - t^-$. Therefore the top of (22) yields the claimed conclusion.

For any given $t \in \mathbb{R}$ we can then find guaranteed bounds for the eigenvalues adjacent to t by solving (23), independently of whether t is in a gap of the essential spectrum or not. This approach has a long history and many other interesting developments are possible, see for example [48], [47, Chapter 4] and the references therein.

If the hypotheses of *a)* and *b)* in Theorem 10 are satisfied, then we know that in fact $F_n\left(t + \frac{1}{2\tau^\pm}\right) = \pm\frac{1}{2\tau^\pm}$ and so $t_\infty^\pm = t + \frac{1}{2\tau^\pm}$. Therefore (23) gives the optimal parameters in the context of Proposition 8. See [21, Theorem 11].

n	$\lambda_{\text{low}}^{\text{up}}$
8	$-0.61^{7975654756025}_{8037239681489}$
10	$-0.61803^{1858947202}_{4107439849}$
12	$-0.6180339^{13904800}_{92920872}$
14	$-0.61803398^{6188916}_{8892611}$
16	$-0.618033988^{663973}_{754684}$
18	$-0.6180339887^{47055}_{50049}$
20	$-0.618033988749^{02}_{99}$

$$\lambda = \frac{1-\sqrt{5}}{2} \approx -0.618033988749895$$

Figure 5. See Example 5. Upper and lower bounds for the eigenvalue λ which is inside the gap of the essential spectrum of the operator B from Example 3. Here the trial spaces $\mathcal{L}_n$ are also as in Example 3.

Remark 4. *The conclusions in a) and b) of Theorem 10 are optimal, only when $t \notin \operatorname{spec} A$. If t is an isolated eigenvalue*[10]*, for example, it is possible to replace these bounds by*

a) $t + \frac{1}{\tau_-} \leq \sup\{\operatorname{spec} A \cap (-\infty, t)\}$

b) $t + \frac{1}{\tau_+} \geq \inf\{\operatorname{spec} A \cap (t, \infty)\}$.

That is, detection of the other points in the spectrum of A which are adjacent to t is always possible. The proof of this assertion can be found in [48]. This observation will have an important consequence when we consider $A = \mathcal{M}$ in §5.4.

Example 5. Let $\mathcal{L}$ and B be as in Example 3. We constructed the table in Figure 5 by finding the largest eigenvalue $\tau^{\text{up}} = \tau_+ > 0$ of (23) with $t = -1$ and the smallest eigenvalue $\tau_{\text{low}} = \tau_- < 0$ of (23) with $t = 1$, then setting

$$\lambda_{\text{low}} = 1 + \frac{1}{\tau_{\text{low}}} \qquad \& \qquad \lambda^{\text{up}} = -1 + \frac{1}{\tau^{\text{up}}} \quad .$$

By Remark 4, we know that $\lambda_{\text{low}} < \lambda < \lambda^{\text{up}}$ where $\lambda = \frac{1-\sqrt{5}}{2}$ is the eigenvalue of B which is inside the gap of the essential spectrum.

A good choice of t is essential in order to ensure the quality of the complementary bound for eigenvalues in the present setting. See [12] for a recent analysis in this direction.

5.3 Poincaré-Friedrichs constants for the gradient

Let $\Omega \subset \mathbb{R}^2$ be an open bounded set. The *homogeneous Poincaré-Friedrichs constant for the gradient* is the smallest $k_{\text{g}} \equiv k_{\text{g}}(\Omega) > 0$ such that

$$\int_\Omega |u|^2 \leq k_{\text{g}} \int_\Omega |\operatorname{grad} u|^2 \qquad \forall u \in H_0^1(\Omega).$$

This inequality is still valid, if we replace k_g by any other upper bound $\tilde{k}_g \geq k_g$. This is an obvious but important observation, as explicit expressions for k_g are only known for very few regions Ω.

[10] In this case $\mathfrak{n}^{\pm}(t) = t$.

Note that $k_g = \frac{1}{c_g}$, where $c_{\rm g}$ is the smallest eigenvalue of the Dirichlet Laplacian in Ω. Recall (5). The Galerkin method does not provide lower bounds for $c_{\rm g}$ directly, but the technique described in §5.2 can be used for that purpose as we describe next.

The self-adjoint operator

$$\underbrace{\begin{bmatrix} 0 & i\,\mathrm{div} \\ i\,\mathrm{grad} & 0 \end{bmatrix}}_{\mathcal{G}} : \underbrace{\begin{matrix} H_0^1(\Omega) \\ \times \\ H(\mathrm{div},\Omega)^2 \end{matrix}}_{\mathrm{dom}\,\mathcal{G}} \longrightarrow \underbrace{\begin{matrix} L^2(\Omega) \\ \times \\ L^2(\Omega)^2 \end{matrix}}_{\mathcal{H}} \tag{24}$$

is strongly indefinite. Moreover,

$$\mathcal{G}\begin{bmatrix} u \\ \underline{\sigma} \end{bmatrix} = \omega \begin{bmatrix} u \\ \underline{\sigma} \end{bmatrix}$$

if and only if $\omega^2 u = -\,\mathrm{div}\,\mathrm{grad}\,u = -\Delta u$. By construction, u always satisfies Dirichlet boundary conditions. Then the spectrum of $\mathcal{G}$ taken to the square power, exactly matches the spectrum of the Dirichlet Laplacian on Ω, except[11] for the extra eigenvalue 0. We are therefore interested in the smallest $\omega^2 > 0$.

In order to find bounds on the minimal $\omega > 0$ for polygons, we can pick $\mathcal{L}_h \subset \mathrm{dom}\,\mathcal{G}$ made of Lagrange elements of order 1 on a mesh $\mathcal{T}_h$,

$$\mathcal{L}_h = \left\{ \begin{bmatrix} u \\ \underline{\sigma} \end{bmatrix} \in [C^0(\overline{\Omega})]^3 : \begin{bmatrix} u\restriction_K \\ \underline{\sigma}\restriction_K \end{bmatrix} \in [\mathbb{P}_1(K)]^3 \ \forall K \in \mathcal{T}_h,\ u\restriction_{\partial\Omega} = 0 \right\}. \tag{25}$$

The parameter t required in Theorem 10 can be found by domain monotonicity, or by more sophisticated means, such as the numerical homotopy method [34, 35]. Let us consider a concrete case of recent interest.

Let $\frac{\pi}{6} \le \phi \le \frac{\pi}{2}$ and

$$l(\phi) = \sqrt{(1-\cos\phi)^2 + (\sin\phi)^2} \quad .$$

The *inner diameter* of the isosceles triangle T_ϕ with vertices at the points

$$(0,0) \qquad (1,0) \qquad (\cos\phi, \sin\phi)$$

is

$$d(\phi) = \max\{|z_1 - z_2| : z_j \in \mathsf{T}_\phi\} = \max\{1, l(\phi)\} \quad .$$

Let Ω_ϕ be the triangle with vertices at the points

$$(0,0) \qquad \left(\frac{1}{d(\phi)}, 0\right) \qquad \left(\frac{\cos\phi}{d(\phi)}, \frac{\sin\phi}{d(\phi)}\right)$$

which has inner diameter equal to 1. Explicit formuli for the Poincaré-Friedrichs constant are known for the particular cases $\phi = \frac{\pi}{3}$ and $\phi = \frac{\pi}{2}$, and a few conjectures [29] exist when $\phi = \frac{\pi}{k}$ for k an integer other than 2 or 3. Therefore this is a natural model to examine numerically. Let us discuss it in some detail.

[11] Note that $u = 0$ and $\underline{\sigma} \neq 0$ is not ruled out for the eigenvalue problem associated to $\mathcal{G}$.

k	$\tilde{c}_{\mathrm{g}}$	k	$\tilde{c}_{\mathrm{g}}$	k	$\tilde{c}_{\mathrm{g}}$
4	104.8235	7	57.8182	10	71.4391
5	80.2329	8	52.6330	11	83.6817
6	66.2661	9	61.2454	12	98.6122

$$\phi = \frac{k\pi}{24}$$

Figure 6. See Example 6. Lower bounds for $c_{\mathrm{g}}(\Omega_\phi)$. The numerical values are found by solving the problem (23) and applying Theorem 10. Here we have fixed $t = 11$ and chosen the trial spaces as in (25) for $h = 0.003$.

The eigenvalues of the Dirichlet Laplacian on $\Omega_{\frac{\pi}{2}}$ can be found by symmetry from those of the square, giving

$$c_g(\Omega_{\frac{\pi}{2}}) = 5\pi^2 d\Big(\frac{\pi}{2}\Big)^2 = 10\pi^2 \approx 98.6960 \quad .$$

The eigenvalues of the Dirichlet Laplacian on the equilateral triangle were computed in the 19th Century, [28, §57]. The smallest eigenvalue for the sides equal to 1 is $\frac{16\pi^2}{3}$, so

$$c_{\mathrm{g}}(\Omega_{\frac{\pi}{3}}) = \frac{16\pi^2 d(\frac{\pi}{3})^2}{3} = \frac{16\pi^2}{3} \approx 52.6379 \quad .$$

It has been recently shown [29] that $\Omega_{\frac{\pi}{3}}$ minimises c_{g} among all other triangles of inner diameter 1. In fact it also minimises the second and third eigenvalue (counting multiplicity), [29, Theorem 1.2].

Example 6. Set

$$t < \frac{4\pi\sqrt{7}}{3} \quad .$$

The right hand side is the square root of the second (and third) eigenvalue on $\Omega_{\frac{\pi}{3}}$, and it is therefore below all other second eigenvalues of the regions Ω_ϕ. By applying Theorem 10-*a)* for that fixed parameter t, we find guaranteed numerical lower bounds for $c_{\mathrm{g}}(\Omega_\phi)$ as illustrated by the table in Figure 6.

5.4 Eigenfrequencies of the resonant cavity

In the Example 2 and the Figure 2, we saw that the Galerkin method is not applicable directly to the operator $\mathcal{M}$, when the trial subspaces were made of standard finite elements. We now show that, by stark contrast, the method describe in §5.2 provides a reliable mean to computing one-side bounds for the non-zero eigenvalues of this operator on these same trial subspaces.

Set $\mathcal{L}_h$ as in (12) and fix $t = 0$. Compute $\tau_+ > 0$ from the reduced eigenvalue problem (23). By virtue of Remark 4-*b)*,

$$\frac{1}{\tau_+} \geq \omega_1,$$

the smallest positive eigenvalue of $\mathcal{M}$. Let us consider experiments with a few concrete regions Ω.

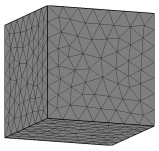
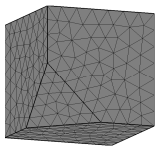
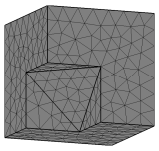
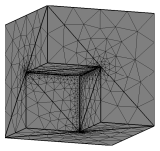
Ω_{c}	Ω_{s}	Ω_{n}	Ω_{F}
			
$\omega_1 \leq 1.414214$	$\omega_1 \leq 1.412218$	$\omega_1 \leq 1.329259$	$\omega_1 \leq 1.142621$
$\dim \mathcal{L}_{0.5} \approx 118\text{K}$	$\dim \mathcal{L}_{0.5} \approx 117\text{K}$	$\dim \mathcal{L}_{0.5} \approx 111\text{K}$	$\dim \mathcal{L}_{0.9} \approx 251\text{K}$

Figure 7. See §5.4 and Figure 11. A numerical approximation of guaranteed upper bounds for the first positive eigenvalue of $\mathcal{M}$ on the region shown. The trial spaces are made of Lagrange elements of order 3 on the corresponding tetrahedral mesh. For the region Ω_{F} we have chosen a fairly large maximal $h = 0.9$, but we made the diameter of the elements substantially smaller near the segments of non-convexity, hence the large dimension of the trial space.

Figure 7 shows numerical approximations of upper bounds for ω_1 on (26)

$$\Omega_{\mathsf{c}} = (0,\pi)^3$$
$$\Omega_{\mathsf{s}} = \Omega_{\mathsf{c}} \setminus \mathsf{T}[(0,0,0);(\pi/2,0,0);(0,\pi/2,0);(0,0,\pi/2)]$$
$$\Omega_{\mathsf{F}} = \Omega_{\mathsf{c}} \setminus [0,\pi/2]^2$$
$$\Omega_{\mathsf{n}} = \Omega_{\mathsf{F}} \cup \mathsf{T}[(\pi/2,\pi/2,\pi/2);(\pi/2,\pi/2,0);(0,\pi/2,\pi/2);(\pi/2,0,\pi/2)]$$

where $\mathsf{T}[p_1;p_2;p_4;p_4]$ is the tetrahedron with vertices p_j. The region Ω_{F} is often called the "Fichera" domain. The trial spaces were constructed on the mesh depicted in each case.

Remark 5. *This approach can also be employed for computing upper bounds of any other positive eigenvalue of $\mathcal{M}$. These upper bounds can be obtained from the subsequent largest positive eigenvalues of* (23)*. Many more results in this respect can be found in [2]. Specifically it has been established in [2] that $\frac{1}{\tau_+} \downarrow \omega_1$ in the regime $h \to 0$. Moreover, whenever Ω is convex, the convergence rate is optimal in a suitable setting.*

6 Global bounds for eigenvalues

The following mechanism for computing eigenvalue bounds does not require an input parameter t. We pick an arbitrary trial subspace of the domain and without any further information, other than the action of A on this trial subspace, it renders rigourous spectral bounds for A. In this respect the technique will be *global* in nature[12]. Recall the question *a)* posed at the beginning of §2.2. The trade-off here is that the generated spectral bounds might not be optimal. Recall the question *b)*.

[12]The Galerkin method is also a global method in this sense.

6.1 The second order spectrum

The remarkable *quadratic method* is known to avoid spectral pollution completely [19, 40, 30] and is convergent [7, 8, 14] to points in the discrete spectrum for any regular family of subspaces.

Fix $\mathcal{L}_n \subset \operatorname{dom} A$. Let $\mathfrak{Q}_n(z)$ and $\tilde{\mathfrak{Q}}_n(z)$ be as in (20). Define

$$G_n(z) = \min_{0 \neq u \in \mathcal{L}_n} \frac{\|\tilde{\mathfrak{Q}}_n(z)u\|}{\|u\|} \qquad \forall z \in \mathbb{C} \quad .$$

According to Lemma 7, $F_n(t)^2 = G_n(t)$ for all $t \in \mathbb{R}$. However, $F_n(z)^2$ and $G_n(z)$ generally differ outside the real axis due to (17). Indeed, note that $G_n(z)$ should always vanish at the zeros of the scalar polynomial $\det \mathfrak{Q}_n(z)$, but this cannot generally be the case for $F_n(z)$. Nonetheless, due to the local regularity of both maps, if $G_n(z)$ is small for z near $\mathbb{R}$, then $F_n(t)$ should also be small for t in a vicinity of $\operatorname{re}(z)$. As we shall see next, spectral enclosures for A can be determined from the zeros of $G_n(z)$ which are close to $\mathbb{R}$. See the figures 4 and 8.

The *second order spectrum of A relative to a trial subspace $\mathcal{L}_n$* is defined as

$$\operatorname{spec}_2(A, \mathcal{L}_n) = \{\zeta \in \mathbb{C} : \det \mathfrak{Q}_n(\zeta) = 0\} \quad .$$

Since $\mathfrak{Q}_n(z)^* = \mathfrak{Q}_n(\overline{z})$, then

$$\operatorname{spec}_2(A, \mathcal{L}_n) = \overline{\operatorname{spec}_2(A, \mathcal{L}_n)} \quad .$$

That is, the points in the second order spectrum form conjugate pairs.

Lemma 11. *If $\zeta \in \operatorname{spec}_2(A, \mathcal{L}_n)$, then $F_n(\operatorname{re}\zeta) \leq |\operatorname{im}\zeta|$.*

Proof. Firstly observe that

$$\det \mathfrak{Q}_n(\zeta) = 0 \quad \iff \quad \exists 0 \neq u \in \mathcal{L}_n,\ \langle \mathfrak{Q}(\zeta)u, v\rangle = 0 \quad \forall v \in \mathcal{L}_n \quad .$$

Let $\zeta = \alpha + i\beta$ for $\alpha, \beta \in \mathbb{R}$. Then

$$\begin{aligned} \langle \mathfrak{Q}(\zeta)u, v\rangle &= \langle (A-\zeta)u, (A-\overline{\zeta})v\rangle \\ &= \langle (A-\alpha)u, (A-\alpha)v\rangle - 2i\beta\langle (A-\alpha)u, v\rangle - \beta^2\langle u, v\rangle \quad . \end{aligned}$$

For $v = u$, it follows from the hypothesis that

$$\|(A-\alpha)u\|^2 - \beta^2\|u\|^2 - 2i\beta\langle (A-\alpha)u, u\rangle = 0 \quad .$$

As the real part of this expression should vanish, then $F_n(\alpha) \leq |\beta|$.

A remarkable universal connection between the second order spectra and the spectrum of any given self-adjoint operator becomes apparent by combining this lemma with (18).

Theorem 12 (Global spectral inclusions)**.** *The following holds true for any self-adjoint operator A and any trial subspace $\mathcal{L}_n \subset \operatorname{dom} A$.*

$$(27)\ \ \zeta \in \operatorname{spec}_2(A, \mathcal{L}_n) \quad \Rightarrow \quad [\operatorname{re}\zeta - |\operatorname{im}\zeta|, \operatorname{re}\zeta + |\operatorname{im}\zeta|] \cap \operatorname{spec} A \neq \varnothing \quad .$$

This statement was first formulated in [40] and its concrete potential for eigenvalue calculation was first highlighted in [30]. Its origins can be traced back to [27].

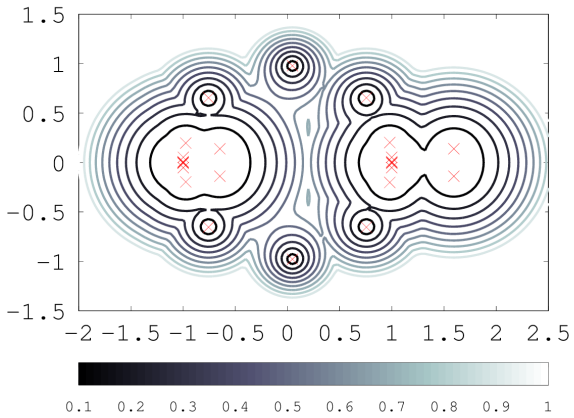


Figure 8. The operator and trial subspace here are as in examples 3, 5 and 7. The contour lines are level sets of $G_{12}(z)$ for $(\operatorname{re} z, \operatorname{im} z) \in [-2, 2.5] \times [-1.5, 1.5]$. The red crosses are $\operatorname{spec}_2(B, \mathcal{L}_{12})$. Compare with Figure 4.

Remark 6. *The original proof of* (27) *given in [40, Lemma 4.1] and [30, Lemma 5.2], involved the following refined version. See also [14, Lemma 2.3]. For $a < b$ denote the open disk with diameter the segment (a, b) by*

$$\mathbb{D}(a, b) = \left\{ z \in \mathbb{C} : \left| z - \frac{a+b}{2} \right| < \frac{b-a}{2} \right\} \quad .$$

Then,

$$(28) \qquad (a, b) \cap \operatorname{spec} A = \varnothing \qquad \Rightarrow \qquad \mathbb{D}(a, b) \cap \operatorname{spec}_2(A, \mathcal{L}_n) = \varnothing \quad .$$

The refined implication (28) yields the following local criterion.

$$(29) \qquad \left. \begin{array}{r} (a, b) \cap \operatorname{spec} A = \{\lambda\} \\ z \in \mathbb{D}(a, b) \cap \operatorname{spec}_2(A, \mathcal{L}_n) \end{array} \right\} \Rightarrow$$
$$\operatorname{re}(z) - \frac{|\operatorname{im}(z)|^2}{b - \operatorname{re}(z)} < \lambda < \operatorname{re}(z) + \frac{|\operatorname{im}(z)|^2}{\operatorname{re}(z) - a} \quad .$$

See [13] and [43]. When the left hand side of (29) is verifiable by analytic means, this gives rise to sharper local eigenvalue bounds, which are some times comparable with those from §5. See [15] and [43].

The segment in (29) will have a smaller length than that in (27), only when suitable $z \in \operatorname{spec}_2(A, \mathcal{L}_n)$ is fairly close to the real line. For regular families of trial subspaces, this is ensured under certain conditions.

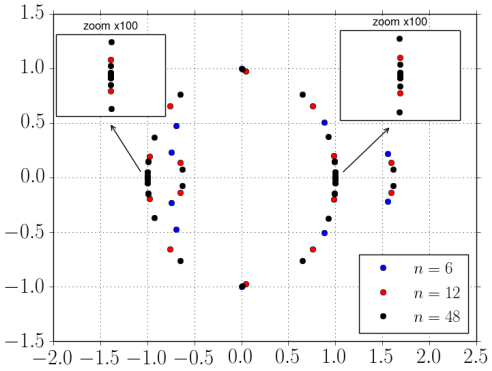
Theorem 13 (Exactness of the quadratic method)**.** *Let $\{\mathcal{L}_n\}_{n \in \mathbb{N}}$ be a regular family of trial subspaces $\mathcal{L}_n \subset \operatorname{dom} A$. For any isolated eigenvalue $\lambda \in \operatorname{spec} A$, there exists $\zeta_n \in \operatorname{spec}_2(A, \mathcal{L}_n)$ such that $|\zeta_n - \lambda| \to 0$.*

For proofs of this statement and further convergence properties of second order spectra see [7, 8, 14, 44, 11].

$$\lambda = \frac{1-\sqrt{5}}{2} \approx -0.618$$

n	(27)	(29)
8	$-0.^{501}_{822}$	$-0.^{586}_{678}$
10	$-0.^{507}_{845}$	$-0.^{587}_{693}$
12	$-0.^{509}_{786}$	$-0.^{593}_{660}$
14	$-0.^{517}_{797}$	$-0.^{599}_{669}$
16	$-0.^{518}_{764}$	$-0.^{598}_{651}$
18	$-0.^{525}_{769}$	-0.6^{05}_{56}
20	$-0.^{525}_{749}$	-0.6^{02}_{45}

(a) Upper and lower bounds for the eigenvalue λ which is inside the gap of the essential spectrum of B as n increases



(b) $\mathrm{spec}_2(B, \mathcal{L}_n)$ for $n \in \{6, 12, 48\}$

Figure 9. See Example 7. As n increases there are conjugate pairs in $\mathrm{spec}(B, \mathcal{L}_n)$ approaching $\mathrm{spec}\, A$. See Theorem 13. Compare the table in (a) with that in Figure 5.

Example 7. Let S be the operator (6) and $\mathcal{L}$ be the regular family (8). Let $K : L^2(-\pi, \pi) \longrightarrow L^2(-\pi, \pi)$ be any other compact operator. Then

$$\begin{aligned} \liminf_{n\to\infty} \mathrm{spec}_2(A+K, \mathcal{L}_n) &= \{\zeta \in \mathbb{C} : \exists \zeta_n \in \mathrm{spec}_2(A+K, \mathcal{L}_n),\ \zeta_n \to \zeta\} \\ &= \{|z| = 1\} \cup \mathrm{spec}_{\mathrm{dsc}}(S+K) \quad . \end{aligned}$$

See [7, Proposition 3 and Lemma 4]. Let $\mathcal{L}$ and B be the same as in examples 3 and 5. In Figure 8 we depict $\mathrm{spec}_2(B, \mathcal{L}_{12})$ alongside with contour lines of $G_{12}(z)$. On the right side of Figure 9 we depict $\mathrm{spec}_2(B, \mathcal{L}_n)$ for $n \in \{6, 12, 48\}$.

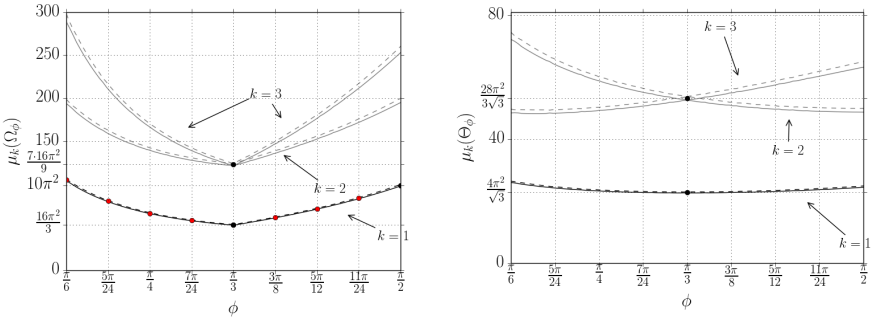
Remark 7. *Convergence of points in the second order spectrum to isolated eigenvalues, only require the angle between* $\ker(A - \lambda)$ *and* $\mathcal{L}_n$ *to be small in the regime* $n \to \infty$. *The rate of convergence* $\zeta_n \to \lambda$ *can be found explicitly in terms of this angle. See [8, 14].*

Let $\mathcal{L}$ and B be the same as in examples 3 and 5. The table in Figure 9-(a) shows enclosures for $\lambda \approx -0.618$. This experiment is similar to that performed in Example 5. Observe that, for the same trial subspaces, the intervals corresponding to (27) are clearly several orders of magnitude larger than those from the table in Figure 5. However, the remarkable fact to be highlighted here is that the former were found *without any preliminary knowledge about* $\mathrm{spec}\, A$.

By virtue of (29), if $\zeta \in \mathrm{spec}_2(B, \mathcal{L}_n) \cap \mathbb{D}(-1, 1)$, then

$$\mathrm{re}(\zeta) - \frac{|\,\mathrm{im}(\zeta)|^2}{1 - \mathrm{re}(\zeta)} < \lambda < \mathrm{re}(\zeta) + \frac{|\,\mathrm{im}(\zeta)|^2}{\mathrm{re}(\zeta) + 1} \quad .$$

The third column of the table in Figure 9-(a) was found by appealing to this observation. The enclosures in the table from Figure 5 are also far



(a) Constant inner diameter **(b)** Constant area

Figure 10. Upper (dashed line) and lower (filled line) bounds for the first three eigenvalues $\mu_k(\Omega_\phi)$ and $\mu_k(\Theta_\phi)$, showing singularities at $\phi = \frac{\pi}{3}$ as appropriate. The numerical values of these bounds were found on 100 uniformly distributed $\phi \in [\frac{\pi}{6}, \frac{\pi}{2}]$, by computing the three conjugate pairs in $\mathrm{spec}_2(\mathcal{G}, \mathcal{L}_{0.002})$ which have minimal positive real part. The red dots show the lower bounds for $c_g(\Omega_\phi)$ from Figure 6 which were calculated by a different method.

smaller than those in this column. As it turns, the local method described in §5.2 beats in this case by several orders of magnitude this other version of a local method.

6.2 The eigenvalues of the Laplacian on triangles

Upper and lower bounds for the eigenvalues of the Laplacian subject to Dirichlet boundary conditions on two-dimensional regions can be found by computing a few points in $\mathrm{spec}_2(\mathcal{G}, \mathcal{L}_h)$, where $\mathcal{G}$ is the operator in (24) and $\mathcal{L}_h$ is a trial subspace as in (25).

As mentioned in Example 6, the equilateral triangle minimises the first three eigenvalues $\mu_1(\Omega) < \mu_2(\Omega) \leq \mu_3(\Omega)$ among any other triangle $\Omega \subset \mathbb{R}^2$ of the same inner diameter. Let us now examine this observation by means of a numerical test.

For $\frac{\pi}{6} \leq \phi \leq \frac{\pi}{2}$, let $\Theta_\phi \subset \mathbb{R}^2$ be the triangle with vertices at the points

$$(0,0) \qquad (2\csc\phi, 0) \qquad (2\cot\phi, 2)$$

which has area equal to 1. It is currently known that $\mu_1(\Theta_{\frac{\pi}{3}})$ is minimal, but not so $\mu_k(\Theta_{\frac{\pi}{3}})$ for $k = 2, 3$. See [41, Theorem 1.3].

Example 8. In Figure 10-(a) we show upper and lower bounds for $\mu_k(\Omega_\phi)$ for 100 uniformly distributed $\phi \in [\frac{\pi}{6}, \frac{\pi}{2}]$. These bounds, consequence of (27), were found by computing the conjugate pairs of the second order spectra which have smaller positive real part. Similarly Figure 10-(b) shows upper and lower bounds for $\mu_k(\Theta_\phi)$ for the same range of ϕ. These graphs can be compared with those from [29, Figures 5 and 6] where only upper bounds for the eigenvalues were displayed.

The Galerkin method can be used for computing upper bounds for μ_k. These upper bounds are far more accurate than the ones determined by the quadratic method in this example. Therefore, when applicable, the Galerkin method is certainly preferable.

The local methods from §5.2 can also be employed for computing complementary bounds for the eigenvalues of L. This is achieved by picking $\mu_k < t < \mu_{k+1}$, for example, then determining a lower bounds for μ_k from Theorem 10-*a)*. Although this is true, and the local bounds turn out to be more accurate, the problem here is how to make sure that $t < \mu_{k+1}$ without knowing the exact value of μ_{k+1}. In some, but not all cases, it is possible to pursue this direction by either appealing to domain monotonicity or by means of numerical homotopy methods [34, 35].

A basic computer code for finding upper and lower bounds for the eigenvalue of the Laplacian by means of the second order spectrum on triangles can be found in §A.

6.3 Eigenfrequencies of the resonant cavity

Let us now revisit the resonant cavity problem (9). Recall Example 2, Figure 2 and §5.4. Set $\mathcal{L}_h$ as in (12). Remarkably, sharp certified information about $\operatorname{spec}\mathcal{M}$ can be extracted from $\operatorname{spec}_2(\mathcal{M}, \mathcal{L}_h)$.

The table in Figure 11 shows computation of spectral enclosures for the regions (26) where the meshes are exactly the same as in Figure 7. These enclosures are consequence of (27). They correspond to the first five conjugate pairs in the second order spectrum, which are closer to the origin and which lie in $\mathbb{D}(0, b)$. Here the parameter b is as given in the table. These are eigenvalue enclosures presumably for ω_m where $m \in \{1, \ldots, 5\}$, but this cannot be guaranteed rigourously at present.

Remark 8. *It is important to realise that here it can not be claimed immediately that there is a single eigenvalue in each segment given in the table in Figure 11. For example note the cluster for $m \in \{2, 3\}$ and $m \in \{4, 5\}$ in the case of Ω_{F}. What is certain, nonetheless, is that each one of these segments intersects* $\operatorname{spec}\mathcal{M}$*. This is already a remarkable fact. Analogously, it cannot be derived from the table in Figure 7 that ω_m for $m > 1$ is above the bound shown in that table. However, it is certain that ω_1 is below that bound. This issue is also present in standard implementations of the Galerkin method.*

7 Further reading

Canonical references on eigenvalue computation for self-adjoint operators include [5], [17, p.283-286], [47], [42, Chapter 6], and the extensive lists of references therein.

Rigourous studies of the phenomenon of spectral pollution along the lines of the discussion presented in §4 can be found in [45], [10], [31], [1] and [37]. A recent survey in the context of Mathematical Physics and Chemistry with a complete bibliography, can be found in [32].

It is quite difficult to trace back the origins of the local method discussed in §5.2, but [47, Chapter 4] has some information on that. In

	Ω_{c}	Ω_{s}	Ω_{n}	Ω_{F}
m	$b = 1.8$	$b = 1.8$	$b = 2.1$	$b = 2.2$
1	1.41^{47}_{37}	1.4^{183}_{061}	$1.^{4402}_{2021}$	$1.^{2554}_{0100}$
2	1.41^{47}_{37}	1.4^{347}_{266}	$1.^{5678}_{4967}$	1.5^{845}_{015}
3	1.41^{47}_{37}	1.4^{344}_{269}	$1.^{5664}_{4982}$	1.5^{830}_{032}
4	1.73^{30}_{11}	1.7^{617}_{486}	2.0329 1.8534	2.1722 1.9863
5	1.73^{31}_{10}	1.7^{612}_{491}	2.0363 1.8506	2.1232 2.0394

Figure 11. See §6.3. A numerical approximation of guaranteed intervals of enclosure for the positive spectrum near the origin of $\mathcal{M}$ on the regions considered in §5.4. The trial spaces are made of Lagrange elements of order 3 on the same tetrahedral mesh as shown in Figure 7.

the literature, this approach is often referred-to as the Temple-Lehman-Goerisch method. Generalisation are due to Maelly [23], and Zimmermann and Mertins [48]. These generalisations are often difficult to implement on practical settings. Some times the remarkable homotopy method, [35] and [34], can be used for that purpose. The topic of optimality of this local method is examined in [12], [21], [20] and [19]. Concrete implementations include computations of bounds for sloshing frequencies [3], the magnetohydrodynamics operator [14] and the Maxwell operator [2].

An increasingly systematic study of concrete implementations of the global method discussed in §6 have been carried out during the last 10 years. These include the one-dimensional elasticity operator [30], perturbed periodic Schrödinger operators [13], the Dirac operator [9], the magnetohydrodynamics operator [43] and complementary bounds for eigenvalues of semi-definite operators [14].

Aside from [26], no systematic comparison of the methods described in §5.2 and §6 has been conducted.

In [45], [46] and [25] a new general method for eigenvalue computation of self-adjoint operators in gaps of the essential spectrum has been proposed. This method uses non-self-adjoint perturbations of the self-adjoint operator in question to "lift" the eigenvalues from the gap of the essential spectrum to the complex plane where spectral pollution does not occur. This method can be traced back to [33] in the case of particular differential operators. It is not impossible that this technique can outperform the two methods described in §5 and §6 in terms of convergence rates. The question is certainly worth exploring in further details.

A Computer codes

Below are two computer codes written in open source software, which combined allow numerical estimation of eigenvalue inclusions for the operator L by means of the second order spectrum. The region Ω is a triangle with vertices $(0, 0)$, $(\pi, 0)$ and $(0, \pi)$, but it can easily be changed to other polygons.

The reduced matrices are generated in FreeFem++ [24] and saved in appropriate files.

```
// LapDBCSecOrd.edp
// FreeFem++ code for assembling the matrices K, L and M
// for the Dirichlet Laplacian
// on  \Omega = triangle with vertices (0,0), (pi,0), (0,pi)
// The files are stored as spec2_Lap_*.dat

border C01(t=0,1){x=t*pi;y=0;label=1;}
border C02(t=0,1){x=(1-t)*pi;y=(t-1)*pi+pi;label=1;}
border C03(t=0,1){x=0;y=(1-t)*pi;label=1;}
int n=30;
mesh Th=buildmesh(C01(n)+C02(n)+C03(n));
plot(Th, wait=false);
fespace Vh(Th,[P1,P1,P1]);
Vh [u,s1,s2],[v,r1,r2];
varf  k([u,s1,s2],[v,r1,r2])=
          int2d(Th)(dx(u)*dx(v)+dy(u)*dy(v)+(dx(s1)+dy(s2))*
                   (dx(r1)+dy(r2)))+on(1,u=0);
varf  l1([u,s1,s2],[v,r1,r2])=
          int2d(Th)(dx(u)*r1+dy(u)*r2+(dx(s1)+dy(s2))*v) ;
varf  m([u,s1,s2],[v,r1,r2])=
          int2d(Th)(u*v+s1*r1+s2*r2);
matrix K= k(Vh,Vh,eps=1e-20);
matrix L1= l1(Vh,Vh,eps=1e-20);
matrix M= m(Vh,Vh,eps=1e-20);
{ofstream file("spec2_Lap_K.dat");  file << K << endl;}
{ofstream file("spec2_Lap_L1.dat");  file << L1 << endl;}
{ofstream file("spec2_Lap_M.dat");  file << M << endl;}
```

The files containing the matrix entries are then read by a script in Octave [22]. Points in the second order spectrum can be computed by means of the following code.

```
# -- Function File: [K,L,M]=LapSpec2(filename1,filename2,filename3)
# [K,L,M] are the matrices extracted from the files filename*
#
# Example:
# >> [K,L,M]=LapSpec2("spec2_Lap_K.dat","spec2_Lap_L1.dat",
# >> ...   "spec2_Lap_M.dat")
# >> S=[eye(size(K)),zeros(size(K));zeros(size(K)),M];
# >> T=[zeros(size(K)),eye(size(K));-K,2*L];
# >> ev=eigs(T,S,6,3);

function [K,L,M]=LapSpec2(filename1,filename2,filename3)
K=getfmat(filename1);
L1=getfmat(filename2);
M=getfmat(filename3);
L=I*triu(L1,1)-I*triu(L1,1)';
```

```
end
# Functions needed to read the files
function A=getfmat(filename)
fid=fopen(filename);
[nn,ix1,ix2,l]=rsm(fid);
fclose(fid);
A=crs(ix1,ix2,l,nn(1));
end
function s=crs(ix1,ix2,a,n1)
s=sparse(ix1,ix2,a,n1,n1);
end
function [nn,ix1,ix2,a]=rsm(fid)
fgetl(fid);
fgetl(fid);
fgetl(fid);
[n1,n2,n3,n4]=fscanf(fid,"%d %d %d %d","C");
nn=[n1,n2];
ix1=zeros(n4,1);
ix2=zeros(n4,1);
a=zeros(n4,1);
for i=1:n4
[ix1(i,1),ix2(i,1),a(i,1)]=fscanf(fid,"%d %d %lf","C");
endfor
end
```

References

[1] D. Arnold, R. Falk, R. Winther, Finite element exterior calculus: from Hodge theory to numerical stability, Bull. Amer. Math. Soc. 47 (2) (2010) 281–354.

[2] G. Barrenechea, L. Boulton, N. Boussaïd, Finite element eigenvalue enclosures for the Maxwell operator, SIAM Journal on Scientific Computing 36 (2014) 2887–2906.

[3] H. Behnke, Lower and upper bounds for sloshing frequencies, Inequalities and Applications (2009) 13–22.

[4] M. Birman, M. Solomyak, The self-adjoint Maxwell operator in arbitrary domains, Leningrad Math. J 1 (1) (1990) 99–115.

[5] D. Boffi, Finite element approximation of eigenvalue problems, Acta Numer. 19 (2010) 1–120.

[6] A. Böttcher, B. Silbermann, Introduction to Large Truncated Toeplitz Matrices, Springer Verlag, Berlin, 1999.

[7] L. Boulton, Limiting set of second order spectra, Math. Comp. 75 (255) (2006) 1367–1382.

[8] L. Boulton, Non-variational approximation of discrete eigenvalues of self-adjoint operators, IMA J. Numer. Anal. 27 (1) (2007) 102–121.

[9] L. Boulton, N. Boussaïd, Non-variational computation of the eigenstates of Dirac operators with radially symmetric potentials, LMS J. Comput. Math (2009) 1–30.

[10] L. Boulton, N. Boussaïd, M. Lewin, Generalised Weyl theorems and spectral pollution in the Galerkin method, J. Spectr. Theory 2 (2012) 329–354.

[11] L. Boulton, A. Hobiny, On the convergence of the quadratic method, IMA J Numer. Anal. (2015) DOI: 10.1093/imanum/drv036

[12] L. Boulton, A. Hobiny, On the quality of complementary bounds for eigenvalues, Calcolo. (2014) DOI: 10.1007/s10092-014-0131-y

[13] L. Boulton, M. Levitin, On approximation of the eigenvalues of perturbed periodic Schrödinger operators, J. Phys. A 40 (31) (2007) 9319–9329.

[14] L. Boulton, M. Strauss, On the convergence of second order spectra and multiplicity, Proc. R. Soc. A 467 (2010) 264–284.

[15] L. Boulton, M. Strauss, Eigenvalue enclosures for the MHD operator, BIT Numerical Mathematics 52 (2012) 801–825.

[16] F. Chatelin, Spectral Approximation of Linear Operators, Academic Press, New York, 1983.

[17] P. Ciarlet, The Finite Element Method for Elliptic Problems, Society for Industrial and Applied Mathematics, Philadelphia, 2002.

[18] E. B. Davies, Spectral Theory and Differential Operators, Cambridge University Press, Cambridge, 1995.

[19] E. B. Davies, Spectral enclosures and complex resonances for general self-adjoint operators, LMS J. Comput. Math 1 (1998) 42–74.

[20] E. B. Davies, A hierarchical method for obtaining eigenvalue enclosures, Math. Comp. 69 (232) (2000) 1435–1455.

[21] E. B. Davies, M. Plum, Spectral pollution, IMA J. Numer. Anal. 24 (3) (2004) 417–438.

[22] J. W. Eaton, D. Bateman, S. Hauberg, GNU Octave Version 3.0.1 Manual: A High-level Interactive Language for Numerical Computations, CreateSpace Independent Publishing Platform, 2009.

[23] F. Goerisch, J. Albrecht, The convergence of a new method for calculating lower bounds to eigenvalues, in: Equadiff 6 (Brno, 1985), vol. 1192 of Lecture Notes in Math., Springer, Berlin, 1986, pp. 303–308.

[24] F. Hecht, New development in FreeFem++, J. Numer. Math. 20 (3-4) (2012) 251–265.

[25] J. Hinchcliffe, M. Strauss, Spectral enclosure and superconvergence for eigenvalues in gaps, Int. Eq. and Oper. Theo. (2015) DOI: 10.1007/s00020-015-2247-0.

[26] A. Hobiny, Enclosures for the eigenvalues of self-adjoint operators and applications to Schrödinger operators, PhD Thesis, Heriot-Watt University, 2014.

[27] T. Kato, On the upper and lower bounds of eigenvalues, J. Phys. Soc. Japan 4 (1949) 334–339.

[28] G. Lamé, Leçons sur la Théorie Mathématique: de l'Élasticité des Corps Solides, Bureau des Longitudes et de l'Ecole Polytechnique, Paris, 1852.

[29] R. Laugesen, B. Siudeja, Dirichlet eigenvalue sums on triangles are minimal for equilaterals, Comm. Anal. Geom. 19 (5) (2011) 855–885.

[30] M. Levitin, E. Shargorodsky, Spectral pollution and second order relative spectra for self-adjoint operators, INA J. Numer. Anal. 24 (2004) 393–416.

[31] M. Lewin, E. Séré, Spectral pollution and how to avoid it (with applications to Dirac and periodic Schrödinger operators), Proc. Lond. Math. Soc. 100 (2010) 864–900.

[32] M. Lewin, E. Séré, Spurious modes in Dirac calculations and how to avoid them, in: Many-Electron Approaches in Physics, Chemistry and Mathematics, Oxford University Press, Oxford, 2014.

[33] M. Marletta, Neumann-Dirichlet maps and analysis of spectral pollution for non-self-adjoint elliptic PDEs with real essential spectrum, IMA J Numer. Anal. 30 (2010) 917–939.

[34] M. Plum, Eigenvalue inclusions for second-order ordinary differential operators by a numerical homotopy method, ZAMP 41 (1990) 205–226.

[35] M. Plum, Bounds for eigenvalues of second-order ordinary differential operators, ZAMP 42 (1991) 848–863.

[36] A. Pokrzywa, Method of orthogonal projections and approximation of the spectrum of a bounded operator, Studia Math. 65 (1) (1979) 21–29.

[37] J. Rappaz, J. Sanchez Hubert, E. Sanchez Palencia, D. Vassiliev, On spectral pollution in the finite element approximation of thin elastic "membrane" shells, Numer. Math. 75 (4) (1997) 473–500.

[38] M. Reed, B. Simon, Methods of Modern Mathematical Physics, vol. I, Academic Press, San Diego, 1980.

[39] M. Reed, B. Simon, Methods of Modern Mathematical Physics, vol. IV, Academic Press, San Diego, 1980.

[40] E. Shargorodsky, Geometry of higher order relative spectra and projection methods, J. Operator Theory 44 (1) (2000) 43–62.

[41] B. Siudeja, Isoperimetric inequalities for eigenvalues of triangles, Indiana Univ. Math. J. 59 (3) (2010) 1097–1120.

[42] G. Strang, G. Fix, An Analysis of the Finite Element Method, Prentice Hall, London, 1973.

[43] M. Strauss, Quadratic projection methods for approximating the spectrum of self-adjoint operators, IMA J. Numer. Anal. 31 (1) (2011) 40–60.

[44] M. Strauss, The second order spectrum and optimal convergence, Math. Comp 82 (284) (2013) 2305–2325.

[45] M. Strauss, The galerkin method for perturbed self-adjoint operators and applications, J Spec. Theo 4 (1) (2014) 113–151.

[46] M. Strauss, A new approach to spectral approximation, J Func. Anal 267 (8) (2014) 3084–3103.

[47] H. F. Weinberger, Variational Methods for Eigenvalue Approximation, Society for Industrial and Applied Mathematics, Philadelphia, 1974.

[48] S. Zimmermann, U. Mertins, Variational bounds to eigenvalues of self-adjoint eigenvalue problems with arbitrary spectrum, Z. Anal. Anwendungen 14 (2) (1995) 327–345.